

GENERATING CONJUGATE DIRECTIONS FOR ARBITRARY MATRICES BY MATRIX EQUATIONS I.

CS.J. HEGEDŰS

Computing Centre, Central Research Institute for Physics, H – 1525 Budapest, POB 49, Hungary

(Received May, 1990)

Abstract. — Recursive methods for generating conjugate directions with respect to an arbitrary matrix are investigated. There are three basic techniques to achieve this aim: (i) minimizing a quadratic form, (ii) generation by projections, and (iii) use of matrix equations. These techniques are equivalent to each other, however, the third one is stressed in this paper because of its versatility. Among matrix equation forms Hestenes – Stiefel type recursions and Lánczos type recursions are mentioned, where the recursion matrices are bidiagonal matrices in the simple case. With respect to the choice of recursion matrices, direct and reverse methods are introduced. The recursion matrices may have lower and upper triangular forms in the direct case and they may be lower and upper Hessenberg matrices in the reverse case. The recursion matrices chosen here are as simple as possible, actually they have no more nonzero elements than that of a bidiagonal matrix. Consequently, the storage of four vectors suffices to perform the recursions in all cases. It is shown that restructuring the bidiagonal matrices makes it possible to avoid zero divisors for the Hestenes – Stiefel type schemes.

1. INTRODUCTION

This is an English language version of three papers that have appeared in Hungarian [1] – [3]. These dealt with conjugate direction methods, and were concerned with generalizing recursions and dealing with the problem of avoiding a zero or nearly zero divisor.

The generalization was carried out on rectangular matrices with the aid of projections, but the resulting recursions in their simple form might break down in the case of zero divisors.

The zero divisor problem was first studied by Luenberger [4], who gave a way to continue the conjugate gradient method if an exact zero divisor occurred for an indefinite symmetric matrix. Later Fletcher [5] did not find the method appropriate because it was not applicable to nearly zero divisors and he searched for more appropriate methods among minimal residual and biconjugate gradient methods.

This introduction gives some basic theorems on conjugate directions with respect to a rectangular matrix A , and the generalized recursions are defined in terms of projections. The resulting recursions then lead to matrix forms which are the starting points of more sophisticated schemes. The resulting schemes are classified as Hestenes – Stiefel and Lánczos type schemes and one can introduce among them *direct* and *reverse* schemes. Although they are briefly considered here, Lánczos type schemes are not dealt with in detail.

The zero divisor – free direct scheme of Hestenes – Stiefel type is given in the second section, while the zero divisor – free reverse scheme of Hestenes Stiefel type can be found in the third section. Theorems are given for the properties of the schemes, hence it is possible to neglect the proofs of some statements in the introduction.

Definition 1.1

The vectors $v_j \in C^m$ and $u_k \in C^n$, $j, k = 1, 2, \dots$ are said to form an A – conjugate or A – biorthogonal system with respect to matrix $A \in C^{m,n}$ if the relations

$$v_j^H A u_k = \alpha_j \delta_{j,k}, \quad \alpha_j \neq 0, \quad j = 1, 2, \dots \quad (1.1)$$

hold, where superscript H denotes the conjugate transpose and $\delta_{j,k}$ is the Kronecker symbol.

This definition is an extension of the orthogonality concept. In the case of special systems and matrices, special terms are used for the A - conjugate vector systems as it can be seen in the following table:

	$v_j \neq u_j$	$v_j = u_j, \alpha_j = 1$
$A = I$	orthogonal	orthonormal
$A = A^T \in \mathcal{R}^{n,n}$	A - orthogonal	A - orthonormal
$A = A^H \in \mathcal{C}^{n,n}$	A - orthogonal	A - unitary

Orthogonal vectors can be generated by the Gram - Schmidt orthogonalization process and this process is readily generalized to produce A - conjugate vector pairs.

Assume the vector system $\{v_j, u_j\}_{j=1}^i$ is A - conjugate and introduce the following left and right projectors:

$$P_i^L = I_m - \sum_{j=1}^i \frac{Au_j v_j^H}{v_j^H Au_j}, \quad P_i^R = I_n - \sum_{j=1}^i \frac{u_j v_j^H A}{v_j^H Au_j}, \quad (1.2)$$

where I_n denotes the unit matrix of n th order. Then one can produce a next A - conjugate pair by

Theorem 1.1

Generalized Gram - Schmidt process: Given the A - conjugate system $\{v_j, u_j\}_{j=1}^i$ and two vectors $r_{i+1} \in \mathcal{C}^m$ and $q_{i+1} \in \mathcal{C}^n$ such that

$$r_{i+1}^H P_i^L A P_i^R q_{i+1} \neq 0, \quad (1.3)$$

then the vectors

$$v_{i+1}^H = r_{i+1}^H P_i^L, \quad u_{i+1} = P_i^R q_{i+1} \quad (1.4)$$

add a new A - conjugate pair to the system $\{v_j, u_j\}_{j=1}^i$.

For proof it is enough to check Definition 1.1 with the projectors in (1.2) and exploit the A - conjugate property of the vectors v_j and u_j .

The systems $\{v_j, Au_j\}$ and $\{A^H v_j, u_j\}$ are biorthogonal by the A - conjugate property, hence the left and right vectors in these systems form linearly independent vector sets.

Definition 1.2

An A - conjugate system is said to be full if it has the maximal number of A - conjugate vector pairs.

Theorem 1.2

The maximal number of A - conjugate pairs is equal to ρ , the rank of matrix A . In the case of a full A - conjugate system, the vectors $A^H v_j, Au_j$ yield a rank - factorization of matrix A .

PROOF. It is seen from Theorem 1.1 and condition (1.3) that it is possible to find a new A - conjugate pair unless $P_i^L A P_i^R = 0$ holds. If this zero condition is written out explicitly by (1.2), one has

$$A = \sum_{j=1}^{\rho} \frac{Au_j v_j^H A}{v_j^H Au_j}, \quad (1.5)$$

which is rank factorization of matrix A , hence the theorem is proven.

By introducing the matrices

$$V = [v_1 \ v_2 \ \dots \ v_\ell], \quad U = [u_1 \ u_2 \ \dots \ u_\ell], \quad (1.6)$$

Eqn. (1.5) can be written in the matrix form:

$$A = AU(V^H AU)^{-1}V^H A, \quad (1.7)$$

where $V^H AU$ is a diagonal matrix. The matrix

$$X = U(V^H AU)^{-1}V^H \quad (1.8)$$

is a (1, 2) - generalized inverse to A because $AXA = A$ is equivalent to (1.7) and $XAX = X$ also holds.

With a full A - conjugate system, it is possible to solve the linear equation $Ax = b$.

Theorem 1.3

If $A \in C^{m,n}$, $\varrho(A) = \varrho, \{v_j, u_j\}_{j=1}^\varrho$ is a full A - conjugate system, $x_1 \in C^n, b \in C^m$ and $r_1 = b - Ax_1$, then the linear system

$$Ax = b \quad (1.9)$$

is consistent if and only if $P_\varrho^t r_1 = 0$. In that case the solutions are given by

$$x = x_1 + \sum_{j=1}^\varrho u_j \frac{v_j^H r_1}{v_j^H A u_j} + P_\varrho^r t, \quad (1.10)$$

where $t \in C^n$ is an arbitrary vector.

PROOF. From fullness of the A - conjugate system $P_\varrho^t A = 0$ follows, cf. (1.5), hence $P_\varrho^t Ax = P_\varrho^t b = 0$ and $P_\varrho^t r_1 = P_\varrho^t (b - Ax_1) = 0$ are necessary consequences. But it is also sufficient because writing out $P_\varrho^t r_1 = 0$ explicitly by (1.2), one has

$$r_1 = b - Ax_1 = A \sum_{j=1}^\varrho u_j \frac{v_j^H r_1}{v_j^H A u_j}, \quad (1.11)$$

and this is reordered into

$$b = A \left[x_1 + \sum_{j=1}^\varrho u_j \frac{v_j^H r_1}{v_j^H A u_j} \right], \quad (1.12)$$

giving a particular solution to (1.9). The general solution is given by (1.10), because $P_\varrho^r t$ is the general solution to the homogeneous linear system $Ax = 0$. Q.E.D.

Simple recursive generation of A - conjugate pairs can be done as follows: Starting with two vectors $r_0 \in C^m$ and $q_0 \in C^n$, compute for $i = 0, 1, \dots$

$$r_{i+1} = P_i^t r_i, \quad q_{i+1}^H = q_i^H P_i^r, \quad (1.13)$$

$$v_{i+1}^H = \varphi_{i+1} r_{i+1}^H C P_i^t, \quad u_{i+1} = \lambda_{i+1} P_i^r K q_{i+1}, \quad (1.14)$$

where $P_0^t = I_m$, $P_0^r = I_n$, λ_{i+1} and φ_{i+1} are freely chosen parameters and the positive definite Hermitian matrices $C \in C^{m,m}$ and $K \in C^{n,n}$ may serve as preconditioners. As will be shown later, vectors r_j form a C - orthogonal system and vectors q_j form a K - orthogonal system. Moreover, adding the A - conjugate property, one can simplify recursions (1.13) - (1.14) to

$$r_{i+1} = r_i - \varphi_i A u_i \frac{\|r_i\|_C^2}{v_i^H A u_i}, \quad q_{i+1}^H = q_i^H - \lambda_i v_i^H A \frac{\|q_i\|_K^2}{v_i^H A u_i} \quad (1.15)$$

$$v_{i+1}^H = \varphi_{i+1} r_{i+1}^H C + \frac{\varphi_{i+1}}{\varphi_i} \frac{\|r_{i+1}\|_C^2}{\|r_i\|_C^2} v_i^H, \quad u_{i+1} = \lambda_{i+1} K q_i + \frac{\lambda_{i+1}}{\lambda_i} \frac{\|q_{i+1}\|_K^2}{\|q_i\|_K^2} u_i, \quad (1.16)$$

where $\|r_i\|_C^2 = r_i^H C r_i$ is the energetic norm of vector r_i and vectors v_0 and u_0 are thought to be zero vectors.

If A is a Hermitian positive definite matrix, then one may take $r_i = q_i$, $v_i = u_i$ for all i and the formulae reduce to the classical Hestenes – Stiefel recursions [6] in the case of $C = K = I$ and to the Hestenes scheme [7], if $C = K \neq I$.

It is worth noting that there are some other basic projection methods for generating A – conjugate directions. Here we mention two others.

In the second scheme, the C – and K – orthogonal projections are introduced by

$$P_i^R = I_m - \sum_{j=1}^i \frac{r_j r_j^H C}{r_j^H C r_j}, \quad P_i^Q = I_n - \sum_{j=1}^i \frac{K q_j q_j^H}{q_j^H K q_j} \quad (1.17)$$

and the recursion formulae are: Choose vectors v_1 , u_1 and for $i = 1, 2, \dots$ compute

$$\lambda_i r_i = P_{i-1}^R A u_i / v_i^H A u_i, \quad \varphi_i q_i^H = v_i^H A P_{i-1}^Q / v_i^H A u_i, \quad (1.18)$$

$$v_{i+1}^H = \lambda_i v_i^H P_i^R, \quad u_{i+1} = \varphi_i P_i^Q u_i, \quad (1.19)$$

where $P_0^R = I_m$ and $P_0^Q = I_n$. These will reduce to the following formulae if $r_0 = 0$ and $q_0 = 0$ are taken by definition:

$$\lambda_i r_i = r_{i-1} + A u_i / v_i^H A u_i, \quad \varphi_i q_i^H = q_{i-1}^H + v_i^H A / v_i^H A u_i, \quad (1.20)$$

$$v_{i+1}^H = \lambda_i v_i^H - r_i^H C / \|r_i\|^2, \quad u_{i+1} = \varphi_i u_i - K q_i / \|q_i\|^2. \quad (1.21)$$

If the scaling parameters λ_i and φ_i are chosen so that $\|r_i\|_C^2 = 1$ and $\|q_i\|_K^2 = 1$, then these formulae get even simpler.

While the previous scheme in its simple form is widely known and applied, this second scheme seems to be unnoticed as yet.

However, they are not basically different from each other. The second scheme (1.20) – (1.21) produces the same A – conjugate directions if the starting pair $\{v_1, u_1\}$ coincides with that of the first method (1.15) – (1.16). They also generate the same r_j, q_j directions if r_1 and q_1 are the same for both methods. The term 'direction' is used because the length of the vectors in the two methods may be different. And switching from one method to the other, they will produce the already generated directions in reverse order. Due to the orthogonal projections and the reversion property, we have called the latter method *reverse* or *orthogonal* recursion.

The third scheme is as follows: For starting vectors $r_0 \in \mathbb{C}^m$ and $q_0 \in \mathbb{C}^n$, compute for $j = 0, 1, \dots$

$$r_{i+1} = P_i^L r_i, \quad q_{i+1}^H = q_i^H P_i^r, \quad (1.22)$$

$$v_{i+1}^H = \varphi_{i+1} q_{i+1}^H B P_i^L, \quad u_{i+1} = \lambda_{i+1} P_i^r B r_{i+1}, \quad (1.23)$$

where matrix $B \in \mathbb{C}^{n,m}$ takes over the role of a preconditioner. Here vectors r_j, q_j form a B – conjugate system and the recursion reduces to

$$r_{i+1} = r_i - A u_i \frac{\varphi_i q_i^H B r_i}{v_i^H A u_i}, \quad q_{i+1}^H = q_i^H - v_i^H A \frac{\lambda_i q_i^H B r_i}{v_i^H A u_i}, \quad (1.24)$$

$$v_{i+1} = \varphi_{i+1} q_{i+1}^H B + v_i^H \frac{\varphi_{i+1}}{\varphi_i} \frac{q_{i+1}^H B r_{i+1}}{q_i^H B r_i}, \quad u_{i+1} = \lambda_{i+1} B r_{i+1} + u_i \frac{\lambda_{i+1}}{\lambda_i} \frac{q_{i+1}^H B r_{i+1}}{q_i^H B r_i}. \quad (1.25)$$

If matrix A is quadratic and B is unit matrix, then this scheme can be shown to be equivalent to the Lánczos tridiagonalization method.

Because of classical analogies, the former schemes will be called Hestenes – Stiefel type and the third scheme will be said to be of Lánczos type. Hestenes – Stiefel type schemes are related to the singular values of matrix $C_1^H A K_1$, where $C = C_1 C_1^H$ and $K = K_1 K_1^H$, while Lánczos

schemes are related to the eigenvalues of BA or AB [1]. These statements can easily be shown by the matrix forms of the recursions which will be introduced here.

In order to do this, collect vectors r_j and q_j into matrices R and Q similarly as was done in (1.6) for vectors v_j , moreover, introduce the diagonal matrices

$$D_a = V^H A U, \quad D_r = R^H C R, \quad D_q = Q^H K Q, \quad (1.26)$$

then the *Hestenes - Stiefel type recursions* can be given by the following matrix equations:

$$LV^H = D_r^{-1} R^H C, \quad A U D_a^{-1} = R L, \quad (1.27)$$

$$D_a^{-1} V^H A = F Q^H, \quad U F = K Q D_q^{-1}, \quad (1.28)$$

where L and F are called *recursion matrices*. For the first direct method L is a lower bidiagonal, F is an upper bidiagonal matrix. On the other hand, L is an upper bidiagonal and F is a lower bidiagonal matrix for the second reverse scheme.

The *Lánczos type schemes* can be given by

$$LV^H = D^{-1} Q^H B, \quad A U D_a^{-1} = R L, \quad (1.29)$$

$$D_a^{-1} V^H A = F Q^H, \quad U F = B R D^{-1}, \quad (1.30)$$

where

$$D = Q^H B R. \quad (1.31)$$

Here the recursion matrices may also have the same forms as before: L is a lower, F is an upper bidiagonal matrix in the 'direct' case, but L is an upper, F is a lower bidiagonal matrix in the 'reverse' case.

The matrix forms are convenient to show how preconditioning techniques can be applied. Assume $A_1 \in \mathbb{C}^{m,m}$ and $A_2 \in \mathbb{C}^{n,n}$ are nonsingular matrices such that

$$A_1 A A_2 \approx \begin{pmatrix} I_\ell & 0 \\ 0 & 0 \end{pmatrix} = E_\ell \in \mathbb{R}^{m,n}, \quad (1.32)$$

where matrices A_1 and A_2 come from an incomplete factorization of matrix A . They are composed of products of unit - matrix plus rank - one - matrix type forms which are applied in LU - decompositions.

Then the preconditioning matrices for the Hestenes - Stiefel type schemes are chosen as

$$C = A_1^H A_1, \quad K = A_2^H A_2. \quad (1.33)$$

With these the recursions (1.15) - (1.16) will need a good amount of work, but there is an easy transition to a simpler form. Introduce the new matrices

$$\hat{R} = A_1 R, \quad \hat{Q} = A_2^H Q, \quad (1.34)$$

$$\hat{V}^H = V^H A_1^{-1}, \quad \hat{U} = \hat{A}_2^{-1} U \quad (1.35)$$

and

$$\hat{A} = A_1 A A_2. \quad (1.36)$$

Then (1.27) - (1.28) can be rewritten into a 'precondition - free' form:

$$L \hat{V}^H = \hat{D}_r^{-1} \hat{R}^H, \quad \hat{A} \hat{U} \hat{D}_a^{-1} = \hat{R} L, \quad (1.37)$$

$$\hat{D}_a^{-1} \hat{V}^H \hat{A} = F \hat{Q}^H, \quad \hat{U} F = \hat{Q} \hat{D}_q^{-1}, \quad (1.38)$$

where the recursion matrices are the same and the diagonal matrices are also unchanged:

$$\hat{D}_r = D_r, \quad \hat{D}_q = D_q, \quad \hat{D}_a = D_a. \quad (1.39)$$

Thus preconditioning matrices need to be applied in determining the first vectors \hat{r}_1 , \hat{q}_1 , \hat{v}_1 and \hat{u}_1 by (1.34) – (1.35) and it is enough to apply only $\hat{A} = A_1 A A_2$ at subsequent steps.

The transition is the same for the Lánczos type scheme and that corresponds to the preconditioning matrix

$$B = A_2 E_q^T A_1, \quad (1.40)$$

where E_q comes from (1.32). Now $\hat{D} = D$, but a simpler 'preconditioner' $\hat{B} = E_q^T$ remains in the formulae.

The linear equation $Ax = b$ may be replaced by

$$\hat{A}\hat{x} = \hat{b} = A_1 b, \quad (1.41)$$

where the final solution is

$$x = A_2^{-1} \hat{x}. \quad (1.42)$$

The solution process can be done by Theorem 1.3. For a starting vector \hat{x}_1 compute $\hat{r}_1 = \hat{b} - \hat{A}\hat{x}_1$ and choose an arbitrary vector \hat{q}_1 to start with the direct schemes. The subsequent approximations are given by (1.10):

$$\hat{x}_{i+1} = \hat{x}_i + \hat{u}_i \frac{\hat{v}_i^H \hat{r}_1}{\hat{v}_i^H \hat{A} \hat{u}_i}. \quad (1.43)$$

One has for the further residual vectors

$$\hat{r}_{i+1} = \hat{b} - \hat{A}\hat{x}_{i+1} = \hat{r}_1 - \sum_{j=1}^i \frac{\hat{A}\hat{u}_j \hat{v}_j^H \hat{r}_1}{\hat{v}_j^H \hat{A} \hat{u}_j} = \hat{P}_i^t \hat{r}_1 = \hat{P}_i^t \hat{r}_1, \quad (1.44)$$

hence the residual vectors coincide with the \hat{r}_j vectors of the recursions as can be seen from (1.13) and (1.22). Vector \hat{r}_1 need not be saved because $\hat{v}_j^H \hat{r}_1$ can be expressed by (1.14) and (1.23) as

$$\hat{v}_j^H \hat{r}_1 = \varphi_j \hat{r}_j^H \hat{P}_{j-1}^t \hat{r}_1 = \varphi_j \|\hat{r}_j\|^2, \quad \|\hat{r}_j\|^2 = \hat{r}_j^H \hat{r}_j \quad (1.45)$$

for the Hestenes – Stiefel scheme and

$$\hat{v}_j^H \hat{r}_1 = \varphi_j q_j^H E_q^T \hat{P}_{j-1}^t \hat{r}_1 = \varphi_j \hat{q}_j^H E_q^T \hat{r}_1 \quad (1.46)$$

for the Lánczos scheme. These quantities would show up in the recursions anyway.

For the reverse scheme, one chooses $\hat{v}_1 = \hat{b} - \hat{A}\hat{x}_1$ and the residual vectors will not coincide with the \hat{r}_j vectors of the recursion. However, the inner product $\hat{v}_j^H \hat{v}_1$ can still be rewritten by (1.19):

$$\hat{v}_j^H \hat{v}_1 = \lambda_{j-1} \hat{v}_{j-1}^H \hat{P}_{j-1}^R \hat{v}_1 = \|\hat{v}_{j-1}\|^2 \prod_{k=1}^{j-2} \lambda_k^{-1} \quad (1.47)$$

for the Hestenes – Stiefel type reverse scheme and vector \hat{u}_1 can be chosen arbitrarily. We do not give here the analogous formula for the reverse Lánczos method, because its projection relations will not be derived in the subsequent sections, although it can be elaborated along similar lines.

As could be seen, there was no need to form the normal equations $A^H A x = A^H b$ to get a positive semidefinite matrix, hence squaring the condition numbers could also be avoided. But the given recursions might break down if $\hat{v}_j^H \hat{A} \hat{u}_j = 0$ at an intermediate stage. Also, a nearly zero divisor might also cause numerical instabilities.

The purpose of these papers is to show that this problem can be cured by restructuring the recursions of Hestenes – Stiefel type. In fact, changing the structure of the recursion matrices L and F is necessary because the nonzero scalars in the recursion matrices may change only the length of the vectors.

2. THE DIRECT RECURSION SCHEME OF HESTENES - STIEFEL TYPE

2.1. The matrix equations

The general matrix recursion equations of Hestenes - Stiefel type are given by (1.26) - (1.28). Now the recursion matrix L is lower bidiagonal in most of the cases but it may change structure into a sparse lower triangular matrix. Also, F is mostly upper bidiagonal, but it may have a sparse upper triangular form. The i -th intermediate stage can be expressed by the equations

$$L_i V^H = D_r^{-1} R^H C, \quad A U D_a^{-1} = R L_i + r_{i+1} \ell_{i+1}^T, \quad (2.1)$$

$$D_a^{-1} V^H A = F_i Q^H + f_{i+1} q_{i+1}^H, \quad U F_i = K Q D_q^{-1}, \quad (2.2)$$

where subscript i in L_i and F_i indicates that i vectors are in the matrices V, U, R and Q . Matrices L_i and F_i are augmented by the following scheme:

$$L_{i+1} = \begin{pmatrix} L_i & 0 \\ \ell_{i+1}^T & 1 \end{pmatrix}, \quad F_{i+1} = \begin{pmatrix} F_i & f_{i+1} \\ 0 & 1 \end{pmatrix}, \quad (2.3)$$

where $\ell_{i+1} \in \mathcal{R}^i$ and $f_{i+1} \in \mathcal{R}^i$.

For simplicity, the diagonal elements are set to 1 in the recursion matrices, hence only the orthogonal r_j and q_j vectors will be scaled.

It follows from the special structure of the recursion matrices that the subsequent vectors can be computed recursively for some r_1, q_1 starting vector pair. One gets from the first equation of (2.1) and from the second equation of (2.2) for $i = 1$:

$$v_1^H = r_1^H C / \|r_1\|_C^2, \quad u_1 = K q_1 / \|q_1\|_K^2. \quad (2.4)$$

If $v_1^H A u_1 \neq 0$ then $\{v_1, u_1\}$ may form the first A -conjugate pair. If $v_1^H A u_1 = 0$ then some of the vectors should be changed. It will be seen that one can always do something if at least one of the vectors $v_i^H A$ and $A u_i$, $i = 1, 2, \dots$ is nonzero.

The possible steps one can take can, in general, be told at the i -th step in general. It will be assumed that the vectors with subscripts $1, 2, \dots, i$ are known; they have the desired orthogonality properties, namely: The system $\{v_j, u_j\}_{j=1}^i$ is A -conjugate, the system $\{r_j\}_{j=1}^i$ is C -orthogonal and the system $\{q_j\}_{j=1}^i$ is K -orthogonal. Further, it is assumed that the i -th order recursion matrices L_i and F_i are known. The generation of the next $(i+1)$ -st vectors can be given by one of the following three steps.

2.2. Bidiagonal step

Here both matrices L and F are continued bidiagonally, hence

$$\ell_{i+1} = -\lambda_{i+1} e_i^T, \quad f_{i+1} = -\varphi_{i+1} e_i. \quad (2.5)$$

Multiply the second equation of (2.1) by e_i from the right and the first equation of (2.2) by e_i^T from the left. As D_a, D_r and D_q are diagonal matrices by the orthogonality assumptions, we have

$$\lambda_{i+1} r_{i+1} = r_i - A u_i / v_i^H A u_i, \quad \varphi_{i+1} q_{i+1} = q_i^H - v_i^H A / v_i^H A u_i. \quad (2.6)$$

Supposing that r_{i+1} and q_{i+1} are nonzero vectors, the numbers λ_{i+1} and φ_{i+1} are freely chosen. For example, they can be set to 1 or such that $\|r_{i+1}\|_C = 1$ and $\|q_{i+1}\|_K = 1$.

For generating the A -conjugate vectors, form L_{i+1} and F_{i+1} by (2.3) and write out the last row of the first equation of (2.1) and the last column of the second equation of (2.2) for subscript $i+1$:

$$v_{i+1}^H = \lambda_{i+1} v_i^H + r_{i+1}^H C / \|r_{i+1}\|_C^2, \quad u_{i+1} = \varphi_{i+1} u_i + K q_{i+1} / \|q_{i+1}\|_K^2. \quad (2.7)$$

If $v_{i+1}^H A u_{i+1} \neq 0$ holds then the vector systems have been successfully extended.

2.3. Diagonal step

This time one of the matrices L and F is augmented diagonally when computing the $(i+1)$ -st vectors. It may be applied only if one of the vectors r_{i+1} and q_{i+1} is exactly zero from the bidiagonal step. Assume q_{i+1} is nonzero, but $\lambda_{i+1}r_{i+1} = r_i - Au_i/v_i^H Au_i = 0$ holds. Then set λ_{i+1} to zero in order that vector r_{i+1} need not be zero. Hence, the border vectors in the recursion matrices are

$$\ell_{i+1}^T = 0, \quad f_{i+1} = -\varphi_{i+1}e_i, \quad (2.8)$$

further assume that

$$\ell_{i+2}^T = 0. \quad (2.9)$$

That is, matrix L is augmented diagonally. As it is related to the left side vector r_{i+1} , it is called a *left diagonal step*.

Substitute matrix L_{i+1} and ℓ_{i+2} of (2.9) into the equations of (2.1). The $(i+1)$ -st row and column yield:

$$r_{i+1} = Au_{i+1}/v_{i+1}^H Au_{i+1}, \quad v_{i+1}^H = r_{i+1}^H C / \|r_{i+1}\|_C^2. \quad (2.10)$$

Due to free scaling, it is possible to give the following solutions to (2.10), these being the formulae of the *left diagonal step*:

$$r_{i+1} = Au_{i+1} / \|Au_{i+1}\|_C, \quad v_{i+1}^H = r_{i+1}^H C, \quad (2.11a)$$

$$r_{i+1} = Au_{i+1}, \quad v_{i+1}^H = r_{i+1}^H C / \|r_{i+1}\|_C^2. \quad (2.11b)$$

They can be checked by substitution. $\|r_{i+1}\|_C = 1$ holds in case a), while $v_{i+1}^H Au_{i+1} = 1$ holds in case b). If $Au_{i+1} = 0$ holds, then $v_{i+1} = 0$ is chosen because it is proportional to CAu_{i+1} , hence the recursion ends in that case.

One can find the formulae of the *right diagonal step* in a similar way:

$$\ell_{i+1}^T = -\lambda_{i+1}e_i^T, \quad f_{i+1} = 0, \quad (2.12)$$

$$q_{i+1}^H = v_{i+1}^H A / \|v_{i+1}^H A\|_K, \quad u_{i+1} = Kq_{i+1}, \quad (2.13a)$$

$$q_{i+1}^H = v_{i+1}^H A, \quad u_{i+1} = Kq_{i+1} / \|q_{i+1}\|_K^2, \quad (2.13b)$$

where $v_{i+1}^H A = 0$ is the end condition. On having one of the choices in (2.13), $f_{i+2} = 0$ will automatically result from (2.6).

2.4. Swap step

Theoretically, this is the most sophisticated step, however, the final result is simple. There are now also *left* and *right side* versions. In the *left step*, one begins with a bidiagonal step at the left side, interchanges the last two orthogonal vectors, say r_i and r_{i+1} , and then recalculates the left conjugate vectors v_i and v_{i+1} . In the derivation below the changing quantities will be indicated by a tilde sign and the final quantities are unsigned. The right vectors q_{i+1} and u_{i+1} have to be calculated last.

The *left swap step* will be elaborated in detail. In order to do this, write the equations of (2.1) into a partitioned form:

$$\begin{pmatrix} L_{i-1} & 0 & 0 \\ \tilde{\ell}_i^T & 1 & 0 \\ 0 & -\tilde{\lambda}_{i+1} & 1 \end{pmatrix} \begin{pmatrix} V_{i-1}^H \\ \tilde{v}_i^H \\ \tilde{v}_{i+1}^H \end{pmatrix} = \begin{pmatrix} D_{r,i-1}^{-1} R_{i-1}^H C \\ \tilde{r}_i^H C / \|\tilde{r}_i\|_C^2 \\ \tilde{r}_{i+1}^H C / \|\tilde{r}_{i+1}\|_C^2 \end{pmatrix}, \quad (2.14)$$

$$AU_i \tilde{D}_a^{-1} = \begin{matrix} R_{i-1} & \tilde{r}_i & \tilde{r}_{i+1} \end{matrix} \begin{pmatrix} L_{i-1} & 0 \\ \tilde{\ell}_i^T & 1 \\ 0 & -\tilde{\lambda}_{i+1} \end{pmatrix} \quad (2.15)$$

where one has $i + 1$ vectors and the subscripts at the matrices indicate previous stages of the recursion steps. Introduce the following matrices

$$P_{i+1} = \begin{pmatrix} I_{i-1} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad G_{i+1} = \begin{pmatrix} I_{i-1} & 0 & 0 \\ 0 & \tilde{\lambda}_{i+1}^{-1} & 1 \\ 0 & 0 & \tilde{\lambda}_{i+1} \end{pmatrix}, \quad Z_i = \begin{pmatrix} I_{i-1} & 0 \\ 0 & -\tilde{\lambda}_{i+1}^{-1} \end{pmatrix}.$$

Now multiply (2.14) from the left by matrix P_{i+1} and write between matrices L and V the product $G_{i+1}G_{i+1}^{-1}$. The result is

$$\begin{pmatrix} L_{i-1} & 0 & 0 \\ 0 & 1 & 0 \\ \tilde{\ell}_i & -\tilde{\lambda}_{i+1} & 1 \end{pmatrix} \begin{pmatrix} V_{i-1}^H \\ -\tilde{\lambda}_{i+1} \tilde{v}_i^H + \tilde{v}_{i+1}^H \\ \tilde{\lambda}_{i+1}^{-1} \tilde{v}_{i+1}^H \end{pmatrix} = \begin{pmatrix} D_{r,i-1}^{-1} R_{i-1}^H \\ \tilde{r}_{i+1}^H / \|\tilde{r}_{i+1}\|_C^2 \\ \tilde{r}_i / \|\tilde{r}_i\|_C^2 \end{pmatrix} C.$$

Similarly, write $P_{i+1}P_{i+1}^{-1}$ between matrices R and L in (2.15) and multiply both sides by Z_i from the right:

$$AU_i \tilde{D}_a^{-1} Z_i = \begin{pmatrix} R_{i-1} & \tilde{r}_{i+1} & \tilde{r}_i \\ 0 & 1 & \\ \tilde{\ell}_i & -\tilde{\lambda}_{i+1}^{-1} & \end{pmatrix}.$$

It can be seen from the last two forms that they have the general form of (2.1) – (2.2), hence after swapping the orthogonal vectors, the new quantities for the *left swap step* are:

$$\ell_i^T = 0, \quad \lambda_{i+1} = \tilde{\lambda}_{i+1}^{-1}, \quad \ell_{i+1}^T = (\tilde{\ell}_i^T, -\lambda_{i+1}), \quad (2.16)$$

$$r_i = \tilde{r}_{i+1}, \quad r_{i+1} = \tilde{r}_i, \quad v_{i+1}^H = \tilde{\lambda}_{i+1}^{-1} \tilde{v}_{i+1}^H, \quad (2.17)$$

$$v_i^H = \tilde{r}_{i+1}^H C / \|\tilde{r}_{i+1}\|_C^2 = -\tilde{\lambda}_{i+1} \tilde{v}_i^H + \tilde{v}_{i+1}^H. \quad (2.18)$$

There is no need to recalculate the inner product $v_i^H A u_i$ because by the A – conjugate property one has

$$v_i^H A u_i = -\tilde{\lambda}_{i+1} \tilde{v}_i^H A u_i + \tilde{v}_{i+1}^H A u_i = -\tilde{\lambda}_{i+1} \tilde{v}_i^H A u_i. \quad (2.19)$$

The right vectors q_{i+1} and u_{i+1} can be calculated only after completing the swap step at the left side.

By observing the tilde sign conventions, the formulae of the *right swap step* are:

$$f_i = 0, \quad \varphi_{i+1} = \tilde{\varphi}_{i+1}^{-1}, \quad f_{i+1}^T = (\tilde{f}_i^T, -\varphi_{i+1}), \quad (2.20)$$

$$q_i = \tilde{q}_{i+1}, \quad q_{i+1} = \tilde{q}_i, \quad u_{i+1} = \tilde{\varphi}_{i+1}^{-1} \tilde{u}_{i+1}, \quad (2.21)$$

$$u_i = K \tilde{q}_{i+1} / \|\tilde{q}_{i+1}\|_K^2 = -\tilde{\varphi}_{i+1} \tilde{u}_i + \tilde{u}_{i+1}. \quad (2.22)$$

$$v_i^H A u_i = -\tilde{\varphi}_{i+1} v_i^H A \tilde{u}_i \quad (2.23)$$

and the $(i + 1)$ – st left vectors need be calculated last with a bidiagonal step.

As can be seen, the i – th and $(i + 1)$ – st orthogonal vectors are swapped and the respective i – th A – conjugate vector is given a value as if it had been restarted by the i – th orthogonal vector, cf. (2.4). After performing swap steps, the border vectors may have more than one nonzero values, because swap steps can be performed successively at one side with increasing indices. For example, if the left vectors are generated successively by swap steps from the k – th to the i – th stage, then matrix L_i will have a 'saw' form:

$$L_i = \begin{pmatrix} \ddots & & & & & \\ \dots & 1 & & & & \\ \dots & -\lambda_{k-1} & 1 & & & \\ & & 0 & 1 & & \\ & & \vdots & \vdots & \ddots & \\ & & \vdots & \vdots & & 1 \\ & & -\lambda_k & -\lambda_{k+1} & \dots & -\lambda_i & 1 \end{pmatrix},$$

showing that the off – diagonal elements $-\lambda_k, -\lambda_{k+1}, \dots$ have been moved to the last row.

2.5. The basic theorem on direct recursion

The following conventions will be introduced. The vectors coming from the bidiagonal steps will be given the tilde sign and they will be considered trial vectors. The accepted vectors will be unsigned. Any vectors are acceptable which add a new element to the corresponding vector sets such that they have the desired properties: Acceptable are vector r_{i+1} if $\|r_{i+1}\|_C \neq 0$, vector q_{i+1} if $\|q_{i+1}\|_K \neq 0$ and the vector pair $\{v_{i+1}, u_{i+1}\}$ if $v_{i+1}^H A u_{i+1} \neq 0$. If $\tilde{r}_{i+1} = 0$ then $\tilde{v}_{i+1} = 0$ is taken and, similarly, if $\tilde{q}_{i+1} = 0$ then vector \tilde{u}_{i+1} is thought to be zero.

Theorem 2.1

If at least one of the vectors $\tilde{v}_{i+1}^H A$ or $A \tilde{u}_{i+1}$ is nonzero, then the vector systems $\{r_j\}_{j=1}^i$, $\{q_j\}_{j=1}^i$ and $\{v_j, u_j\}_{j=1}^i$ can successfully be extended by one of the previously given steps. Specifically,

1. The bidiagonal step is successful, if $\tilde{v}_{i+1}^H A \tilde{u}_{i+1} \neq 0$;
2. The diagonal step is successful
 - on the left side, if $\tilde{r}_{i+1} = 0$ and $A \tilde{u}_{i+1} \neq 0$,
 - on the right side, if $\tilde{q}_{i+1} = 0$ and $\tilde{v}_{i+1}^H A \neq 0$;
3. If $\tilde{v}_{i+1}^H A \neq 0$, then at least one of the bidiagonal or left swap steps is successful.
If $A \tilde{u}_{i+1} \neq 0$, then at least one of the bidiagonal or right swap steps is successful.

2.6. Proof of the basic theorem

There will be some theorems proven which are interesting in themselves.

Theorem 2.2

Vector $r_{i+1} \neq 0$ is C -orthogonal to the vectors r_j , $j = 1, 2, \dots, i$ and vector $q_{i+1} \neq 0$ is K -orthogonal to the vectors q_j , $j = 1, 2, \dots, i$.

PROOF. Assume at first that $\ell_{i+1} \neq 0$ and $f_{i+1} \neq 0$ hold. Multiply the two equations of (2.1) with each other, left side by left side and right side by right side:

$$L_i V^H A U D_a^{-1} = D_r^{-1} R^H C R L_i + D_r^{-1} R^H C r_{i+1} \ell_{i+1}^T.$$

By the definition of matrices D_a and D_r in (1.26), one gets the simple form

$$0 = R^H C r_{i+1} \ell_{i+1}^T.$$

which states for $\ell_{i+1} \neq 0$, that vector r_{i+1} is C -orthogonal to the vectors r_j , $j = 1, 2, \dots, i$. The K -orthogonality of vector q_{i+1} can be shown similarly by multiplying the equations of (2.2).

Observe that conditions $\ell_{i+1} \neq 0$ and $f_{i+1} \neq 0$ are fulfilled if a bidiagonal step or a swap step have been applied. Thus the theorem is yet to be proven for the diagonal step, where $\ell_{i+1} = 0$ or $f_{i+1} = 0$. This case will be shown in the fourth step of proving the next theorem.

Theorem 2.3

The $(i+1)$ -st vectors can be given by the following projections:

$$r_{i+1} = \begin{cases} \lambda_{i+1}^{-1} P_i^\ell r_i, & \text{if } \ell_{i+1} \neq 0, \\ (v_{i+1}^H A u_{i+1})^{-1} \|q_{i+1}\|_K^{-2} P_i^\ell A K q_{i+1}, & \text{if } \ell_{i+1} = 0, \end{cases} \quad (2.24.)$$

$$q_{i+1}^H = \begin{cases} \varphi_{i+1}^{-1} q_i^H P_i^r, & \text{if } f_{i+1} \neq 0, \\ (v_{i+1}^H A u_{i+1})^{-1} \|r_{i+1}\|_C^{-2} r_{i+1}^H C A P_i^r, & \text{if } f_{i+1} = 0, \end{cases} \quad (2.25)$$

$$v_{i+1}^H = \|r_{i+1}\|_C^{-2} r_{i+1}^H C P_i^\ell, \quad (2.26)$$

$$u_{i+1} = \|q_{i+1}\|_K^{-2} P_i^r K q_{i+1}, \quad (2.27)$$

where the projectors P_i^ℓ and P_i^r were introduced in (1.2).

PROOF. It is done in six steps. (i) We begin by deriving some necessary relations. Projector P_i^ℓ can be written in matrix form as

$$P_i^\ell = I_m - AUD_a^{-1}V^H, \quad (2.28)$$

where index i indicates also that matrices V and U have i column vectors. Substitute here AUD_a^{-1} from the second equation of (2.1) and L_iV^H from the first equation. The result is

$$P_i^\ell = I_m - RL_iV^H - r_{i+1}\ell_{i+1}^T V^H = I_m - RD_r^{-1}R^H C - r_{i+1}\ell_{i+1}^T V^H. \quad (2.29)$$

One gets similarly from (2.2)

$$P_i^r = I_n - UD_a^{-1}V^H A = I_n - KQD_q^{-1}Q^H - Uf_{i+1}q_{i+1}^H. \quad (2.30)$$

We get another relation if the first equation of (2.1) is multiplied by matrix R from the right. On applying (1.26) to matrix D_r we obtain that matrices L_i and $V^H R$ are inverses to each other:

$$L_iV^H R = I_i = V^H RL_i. \quad (2.31)$$

(ii) Assume that $\ell_{i+1} \neq 0$ holds. Then $\lambda_{i+1} \neq 0$ follows, which case belongs to the bidiagonal or swap step. Multiply the second equation of (2.1) by vector e_i from the right and exploit (2.31):

$$\lambda_{i+1}r_{i+1} = (RL_i - AUD_a^{-1})e_i = (I_m - AUD_a^{-1}V^H)RL_ie_i = P_i^\ell r_i.$$

The result is the first relation of (2.24). The first relation of (2.25) can be shown similarly on the assumption that $f_{i+1} \neq 0$.

(iii) Assume again $\ell_{i+1} \neq 0$. Multiply (2.29) by vector $r_{i+1}^H C$ from the right:

$$r_{i+1}^H C P_i^\ell = r_{i+1}^H C - \|r_{i+1}\|_C^2 \ell_{i+1}^T V_i^H. \quad (2.32)$$

We have applied here Theorem 2.2, which was shown for the case of $\ell_{i+1} \neq 0$. Now write out the general expression of v_{i+1}^H from the first equation of (2.1). Also substitute here L_{i+1} of (2.3), take the last row and observe that the $(i+1)$ -st order matrix D_r is diagonal:

$$\ell_{i+1}^T V_i^H + v_{i+1}^H = \|r_{i+1}\|_C^{-2} r_{i+1}^H C. \quad (2.33)$$

If this is compared to (2.32), it yields (2.26). The proof of (2.27) is analogous, if $f_{i+1} \neq 0$.

(iv) We finish the proof of Theorem 2.2 for the case of $\ell_{i+1} = 0$, which belongs to the left diagonal step. This time $f_{i+1} \neq 0$ should hold, thus we have (2.27) by step (iii). One can check the identity

$$V_i^H A P_i^r \equiv 0 \quad (2.34)$$

directly by (2.30). From this $V_i^H A u_{i+1} = 0$ follows by (2.27). But the direction of vector r_{i+1} is the same as that of vector $A u_{i+1}$ because of (2.11), hence $V_i^H r_{i+1} = 0$ follows. Finally substitute V_i^H from the first equation of (2.1), and omit the nonsingular matrices L_i and D_r . This leads to

$$R^H C r_{i+1} = 0, \quad R^H \in C^{i,m}, \quad (2.35)$$

which shows the desired orthogonality relations.

The K -orthogonality of vector q_{i+1} can be shown analogously, thus Theorem 2.2 is proven.

(v) By Theorem 2.2 — which is now valid for all cases — repeat step (iii) once again, hence (2.26) and (2.27) are proven for all cases.

(vi) The second formula of (2.24) is really a complicated form of the first equation of (2.11). It has been produced by substituting u_{i+1} of (2.27) into the first equation of (2.11) and the identity

$$A P_i^r \equiv P_i^\ell A \quad (2.36)$$

was made use of. It is necessary to re-write the formula because it shows that the $(i+1)$ -st vectors are generated by projections in every case.

The proof of (2.25) can be complemented similarly, hence the theorem is proven.

Theorem 2.4

If v_{i+1}^H and u_{i+1} are accepted vectors by convention, i.e. $v_{i+1}^H A u_{i+1} \neq 0$, then the vector pair v_{i+1} , u_{i+1} adds a new pair to the system $\{v_j, u_j\}_{j=1}^i$, where $v_{i+1}^H A u_j = 0$ and $v_j^H A u_{i+1} = 0$, $j \leq i$.

PROOF. It is enough to check the projection forms (2.26) and (2.27) by Theorem 1.1.

At this point the orthogonality properties of the generated vectors are shown. We still need to make some comments on swap steps.

If a swap step is made after diagonal steps, then the diagonal recursion matrix will open into a bidiagonal form. Swap steps cannot be made at both sides at the same time because the i -th vectors on the other side are fixed. However, at the i -th stage a right swap step can be made after a left swap step such that they belong to the i -th level, but this time some vectors have to be discarded.

The general forms of the border vectors are

$$\ell_{i+1} = \sum_{j=k}^i -\lambda_{j+1} e_j, \quad f_{i+1} = \sum_{j=h}^i -\varphi_{j+1} e_j, \quad (2.37)$$

where the lower limits of the sums have to fulfil one of the conditions

$$k \leq i, \quad h = i, \quad (2.38a)$$

$$k = i, \quad h \leq i. \quad (2.38b)$$

Theorem 2.5

In the case of the left swap step, one has the relation

$$\varphi_{i+1} \|q_{i+1}\|_K^2 v_{i+1}^H A u_{i+1} = \frac{\|A^H \tilde{v}_{i+1}\|_K^2}{\tilde{\lambda}_{i+1}^2 u_i^H A^H \tilde{v}_i} + \tilde{\lambda}_{i+1}^{-1} \|\tilde{q}_{i+1}\|_K^2 \tilde{\varphi}_{i+1} \tilde{v}_{i+1}^H A \tilde{u}_{i+1} \quad (2.39)$$

and in the case of the right swap step the corresponding formula is

$$\lambda_{i+1} \|r_{i+1}\|_C^2 v_{i+1}^H A u_{i+1} = \frac{\|A \tilde{u}_{i+1}\|_C^2}{\tilde{\varphi}_{i+1}^2 \tilde{u}_i^H A^H \tilde{v}_i} + \tilde{\varphi}_{i+1}^{-1} \|\tilde{r}_{i+1}\|_C^2 \tilde{\lambda}_{i+1} \tilde{v}_{i+1}^H A \tilde{u}_{i+1}. \quad (2.40)$$

PROOF. Only the left swap step formula will be shown. From the second equation of (2.8) and by the A -conjugate property one has

$$v_{i+1}^H A u_{i+1} = v_{i+1}^H A K q_{i+1} / \|q_{i+1}\|_K^2. \quad (2.41)$$

One can substitute here q_{i+1} from (2.7), which is a linear combination of q_i and $A^H v_i$. Vector q_i is multiplied by matrix K in (2.41), but vector $K q_i$ is a linear combination of vectors u_i and u_{i-1} , hence vector q_i yields a zero term because of A -conjugacy:

$$v_{i+1}^H A u_{i+1} = v_{i+1}^H A K A^H v_i \frac{1}{\|q_{i+1}\|_K^2 u_i^H A^H \tilde{v}_i \varphi_{i+1}}.$$

Substitute here vector v_i of (2.18) and vector v_{i+1} of (2.17):

$$v_{i+1}^H A u_{i+1} = -\frac{\|A^H \tilde{v}_{i+1}\|_K^2}{\tilde{\lambda}_{i+1} \varphi_{i+1} \|q_{i+1}\|_K^2 u_i^H A^H \tilde{v}_i} + \frac{\tilde{v}_{i+1}^H A K A^H \tilde{v}_i}{\varphi_{i+1} \|q_{i+1}\|_K^2 u_i^H A^H \tilde{v}_i}. \quad (2.42)$$

Vector $K A^H \tilde{v}_i$ in the second term is the linear combination of vectors u_{i-1} , u_i and \tilde{u}_{i+1} . This can be seen from (2.7), where vector $A^H \tilde{v}_i$ can be expressed by q_i and \tilde{q}_{i+1} and these multiplied by matrix K can be expressed by vectors u_{i-1} , u_i and \tilde{u}_{i+1} according to (2.8). Once again, because of A -conjugacy \tilde{u}_{i+1} is the only vector, which may have a nonzero contribution in the second term. Now, collecting the scalar multipliers from the substitutions, one gets the final formula (2.39), if (2.19) is still considered. Q.E.D.

It can be seen from relation (2.39) that if $\|A^H \tilde{v}_{i+1}\|_K \neq 0$, then one of the inner products $v_{i+1}^H A u_{i+1}$ and $\tilde{v}_{i+1}^H A \tilde{u}_{i+1}$ may be zero. The same is shown by (2.40) for the case of $\|A \tilde{u}_{i+1}\|_C \neq 0$ hence the third statement of Theorem 2.1 follows. The first and second statements of Theorem 2.1 can easily be checked, thus the basic theorem is also proven.

2.7. The inverse of the recursion matrices

When analysing the recursions, we may need the inverses of the recursion matrices L and F . They can be stated by the bordered formula

$$\begin{pmatrix} L_i & 0 \\ \ell_{i+1}^T & 1 \end{pmatrix}^{-1} = \begin{pmatrix} L_i^{-1} & 0 \\ -\ell_{i+1}^T L_i^{-1} & 1 \end{pmatrix}.$$

The inverse can be expressed by a closed formula if the concept of leading and inner elements is introduced for matrix L . Call the first nonzero element of vector ℓ_{i+1}^T a leading element and the subsequent elements up to the diagonal, inner elements. Then introduce the notations

$$\nu_i = \begin{cases} \lambda_i, & \text{if } \lambda_i \text{ is a leading element} \\ 1, & \text{if } \lambda_i \text{ is an inner element,} \end{cases} \quad (2.43)$$

$$\gamma_i = \begin{cases} 1, & \text{if } \ell_i \neq 0, \\ 0, & \text{if } \ell_i = 0, \end{cases} \quad (2.44)$$

$$\beta_i = \begin{cases} 1, & \text{if } \lambda_{i+1} \text{ is a leading element,} \\ \lambda_{i+1}, & \text{if } \lambda_{i+1} \text{ is an inner element.} \end{cases} \quad (2.45)$$

Then the inverse can be expressed by the following formula:

$$(L^{-1})_{ij} = \begin{cases} \gamma_i \beta_j \prod_{k=j+1}^i \nu_k, & i > j, \\ 1, & i = j \\ 0, & i < j. \end{cases} \quad (2.46)$$

The inverse of matrix F^T can be given similarly.

2.8. An example

Consider the following matrix of rank 3:

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{pmatrix},$$

and choose starting vectors $r_1^T = v_1^T = (0 \ 0 \ 1)$ and $q_1^T = u_1^T = (1 \ 0 \ 0)$. Compute the next vectors by bidiagonal steps and choose scaling parameters λ_j and φ_j equal to 1. Then one gets the vectors:

$$r_2 = \frac{1}{3} \begin{pmatrix} -1 \\ -2 \\ 0 \end{pmatrix}, \quad v_2 = \frac{1}{5} \begin{pmatrix} -3 \\ -6 \\ 5 \end{pmatrix}, \quad r_3 = \frac{1}{4} \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix},$$

$$q_2 = \frac{1}{3} \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}, \quad u_2 = \frac{1}{2} \begin{pmatrix} 2 \\ -3 \\ -3 \end{pmatrix}, \quad q_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Vector q_3 turned into 0. If one makes a right diagonal step according to (2.13b), then vectors $q_3^T = v_3^T A = (0 \ 2 \ -2)$, and $u_3^T = (0 \ 1/4 \ -1/4)$ make the systems full. With these vectors the recursion matrices are

$$L = \begin{pmatrix} 1 & & \\ -1 & 1 & \\ & -1 & 1 \end{pmatrix}, \quad F = \begin{pmatrix} 1 & -1 & \\ & 1 & \\ & & 1 \end{pmatrix}. \quad (2.47)$$

Another possibility of avoiding zero vectors is to make a left swap step between the second and third stage. This time vectors r_2 and r_3 are swapped and vectors v_3 and v_2 get new values from (2.17) and (2.18):

$$r_2 = \frac{1}{4} \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \quad v_2 = \frac{4}{5} \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \quad r_3 = \frac{1}{3} \begin{pmatrix} -1 \\ -2 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Thus (2.7) and (2.8) yield the following values for vectors q_3 and u_3 :

$$q_3 = \frac{1}{3} \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 0 \\ 2 \\ -3 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 0 \\ -5 \\ 5 \end{pmatrix}, \quad u_3 = \frac{1}{2} \begin{pmatrix} 2 \\ -3 \\ -3 \end{pmatrix} + \frac{3}{10} \begin{pmatrix} 0 \\ -5 \\ 5 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 5 \\ -9 \\ -6 \end{pmatrix}.$$

Despite our having obtained other vectors, it can still be checked that the generated systems are full. This time, matrices L and F have the following forms:

$$L = \begin{pmatrix} 1 & & \\ & 1 & \\ -1 & -1 & 1 \end{pmatrix}, \quad F = \begin{pmatrix} 1 & -1 & \\ & 1 & -1 \\ & & 1 \end{pmatrix}. \quad (2.48)$$

2.9. Conclusions

A number of ways of generating A - conjugate vectors with respect to an arbitrary nonzero matrix A were studied. Judging from Theorem 1.1 the Gram - Schmidt orthogonalization process can adequately be extended to generate these vectors from a linearly independent base. In fact, Theorem 1.1 is a slight extension to that given by Fox et al. [9] for a symmetric matrix. The maximal number of A - conjugate vector pairs is equal to the rank of matrix A . A full system of A - conjugate vector pairs yields a rank - factorization of matrix A by (1.5) and a (1,2) - generalized inverse to matrix A by (1.8). The complete solution of a consistent linear system in terms of A - conjugate vector pairs is given by Theorem 1.3.

Recursive generation methods result if the base vectors are also generated by projections. Among them Hestenes - Stiefel and Lánczos type methods are given in a preconditioned form. The Hestenes - Stiefel type methods are related to the singular values and the Lánczos type methods are related to the eigenvalues of the respective preconditioned matrix to matrix A . The projective recursion schemes lead to matrix equation forms, where the conjugate vectors and the auxiliary vectors are collected into a matrix. A recursion is characterized by the recursion matrices L and F . Matrix L is lower bidiagonal and matrix F is an upper bidiagonal matrix in the direct schemes. It is also possible to introduce reverse schemes, where matrix L is an upper and matrix F is a lower bidiagonal matrix. However, these methods may break down if a zero or almost zero divisor occurs.

To overcome the problem, the only possible way is to change the structure in the recursion matrices because the nonzero entries in them are only able to influence the length of the vectors. Simple modifications to the direct scheme of Hestenes - Stiefel type are given in Sec. 2, where the recursion matrices may have bidiagonal and diagonal matrix portions together with portions of matrices having a 'sawtooth' form. It is shown that the modified recursion matrices will not spoil the orthogonality properties of the vector systems. The resulting method is equivalent to a projective method and subsequent vectors can always be generated if at least one of Au_i or $A^H v_i$ is a nonzero vector. The equivalence to methods minimizing a quadratic form was shown essentially by Hestenes [7], who proved that any finite method for solving a linear system can be given in terms of minimizing a quadratic form. The resulting recursion matrices in the modified forms are so simple that they have a closed inverse formula.

REFERENCES

1. Cs.J. Hegedűs, L. Bodócs, Generation of conjugate directions: The method of conjugate pairs (In Hungarian), *Alkalmaz. Mat. Lapok* 11, 297-309 (1985).

2. Cs.J. Hegedűs,, General recursions of Hestenes – Stiefel type for conjugate directions. I. The direct recursion (In Hungarian), *Alkalmaz. Mat. Lapok* 13, 83–96 (1987–88).
3. Cs.J. Hegedűs, General recursions of Hestenes – Stiefel type for conjugate directions. II. The reverse recursion (In Hungarian), *Alkalmaz. Mat. Lapok* 13, 361–372 (1987–88).
4. D.G. Luenberger, Hyperbolic pairs in the method of conjugate gradients, *SIAM J. Appl. Math.* 17, 1263–1267 (1969).
5. R. Fletcher, Conjugate gradient methods for indefinite systems, In *Proc. of the Dundee Biennial Conference on Numerical Analysis*; Ed.G.A. Watson Lecture Notes in Mathematics, 506, 73–89 (1976).
6. M.R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Standards Sect. B* 49, 409–436 (1952).
7. M.R. Hestenes, The conjugate gradient method for solving linear systems, In *Proc. Symposia on Applied Mathematics VI*, 83–102 (1956), *Numerical Analysis*, McGraw – Hill, New York, pp., ().
8. C. Lánosz, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Standards Sect. B* 45, 255–283 (1950).
9. L. Fox, H.D. Huskey, J.H. Wilkinson, Notes on the solution of algebraic linear simultaneous equations, *Quart. J. Mech. Appl. Math.* 1, 149–173 (1948).